

AD\_\_\_\_\_

Award Number: W81XWH-11-2-0225

TITLE: Integrative Lifecourse and Genetic Analysis of Military Working Dogs

PRINCIPAL INVESTIGATOR: Carlos Alvarez

CONTRACTING ORGANIZATION: The Ohio State University College of Veterinary  
Medicine

REPORT DATE: October 2012

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command  
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;  
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

<b>REPORT DOCUMENTATION PAGE</b>				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. <b>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</b>					
<b>1. REPORT DATE</b> 24 October 2012		<b>2. REPORT TYPE</b> Annual		<b>3. DATES COVERED</b> 15 September 2011 – 14 September 2012	
<b>4. TITLE AND SUBTITLE</b> Integrative Lifecourse and Genetic Analysis of Military Working Dogs				<b>5a. CONTRACT NUMBER</b>	
				<b>5b. GRANT NUMBER</b> W81XWH-11-2-0225	
				<b>5c. PROGRAM ELEMENT NUMBER</b>	
<b>6. AUTHOR(S)</b> Carlos Alvarez  <b>E-Mail:</b> guillermo.couto@cvm.osu.edu				<b>5d. PROJECT NUMBER</b>	
				<b>5e. TASK NUMBER</b>	
				<b>5f. WORK UNIT NUMBER</b>	
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> Department of Veterinary Clinical Sciences The Ohio State University Columbus, OH 43210				<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>	
<b>9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b> U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b>	
				<b>11. SPONSOR/MONITOR'S REPORT NUMBER(S)</b>	
<b>12. DISTRIBUTION / AVAILABILITY STATEMENT</b> Approved for Public Release; Distribution Unlimited					
<b>13. SUPPLEMENTARY NOTES</b>					
<b>14. ABSTRACT</b>  The purpose of this proposal is to provide insight into gene-environment interactions in a population of military working dogs (MWD). It leverages the simplified genetics and detailed records of the MWD population. There are several critical aspects to meeting the aims of this proposal. 1) development of data-driven selection criteria, 2) biological sampling of representative dogs, and 3) generation of mathematical methodologies capable of handling heterogenous data and statistical tests in consistent manner and providing clear and understandable results that are biologically valid. Each of these criteria posses their own challenges that must be overcome before the project will be successful. We accomplished all of the first year objectives, as described in the body.					
<b>15. SUBJECT TERMS</b> None provided					
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b>  UU	<b>18. NUMBER OF PAGES</b>  7	<b>19a. NAME OF RESPONSIBLE PERSON</b> USAMRMC
<b>a. REPORT</b> U	<b>b. ABSTRACT</b> U	<b>c. THIS PAGE</b> U			<b>19b. TELEPHONE NUMBER</b> (include area code)

## Table of Contents

	<u>Page</u>
<b>Introduction.....</b>	<b>4</b>
<b>Body.....</b>	<b>5-8</b>
<b>Key Research Accomplishments and Reportable Outcomes .....</b>	<b>5-8</b>
<b>Conclusion.....</b>	<b>8</b>

## **Introduction**

The purpose of this proposal is to provide insight into gene-environment interactions in a population of military working dogs (MWD). It leverages the simplified genetics and detailed records of the MWD population. There are several critical aspects to meeting the aims of this proposal. 1) development of data-driven selection criteria, 2) biological sampling of representative dogs, and 3) generation of mathematical methodologies capable of handling heterogeneous data and statistical tests in consistent manner and providing clear and understandable results that are biologically valid. Each of these criteria poses their own challenges that must be overcome before the project will be successful. Here we will provide a breakdown of the previous year's work that has occurred and document our progress towards achieving the specific aims we proposed.

## Body

Task 1- Regulatory Approval: The regulatory approval is currently with Lackland AFB (LAFB). We have received Exempt status for our data protocol and our biological sampling protocol is currently under review. We also have received approval from the Joint Pathology Center to review their records. JPC has also sent copies of all pathology reports associated with the DoD puppy program. My personal involvement in year one has been limited to close personal and electronic/telephone interactions with the Military Veterinary Corp at LAFB. I visited LABD twice, evaluated the medical records system, and discussed future approaches to data analysis with the LAFB veterinarians.

Task 2- Data Capture of Veterinary Records: We have received data from the TSA. This data was used to establish the preliminary database. In addition we have identified software that will be highly efficacious to capture data from the pathology records. We have also established all of the necessary computational infrastructure at Lackland AFB. We have a high-speed scanner in place and a high-performance computer to process all of the data collected. We have also conducted regular conference with Lackland AFB and our site visits have yielded additional collaborations.

Task 3-Methodology Development: Tasks 3a and 3b are complete. Tasks 3c-3h are scheduled for future reporting periods. Task 3g is in progress. We have submitted articles and received comments. We are rewriting the manuscripts to address reviewers concerns. Over the past year we have met multiple relevant scientific and logistic benchmarks we have also encountered and solved several problems associated with the study. Most notably we have two publications (appendix 1 and appendix 2) that were submitted and we are now making the changes recommended by the reviewers. In regard to the statement of work we have met several key benchmarks including development of novel analytic techniques. These techniques (SOW Task 3a-3c) are critical to the success of this project. Specifically we have developed a technique we refer to as Variable Genomic Intersect Analysis or GIA for short. GIA overcomes limitations of traditional genomewide genetic association studies by 1) accounting for latent variables without information loss or distortion, 2) not requiring multiplicity testing correction and 3) significantly increasing power. This method is computationally fast and broadly applicable to simple and complex genetics, as well as to genomics. Although detailed in our manuscript it is important to include here as well. Again, my personal involvement in this area has been limited to interacting with the Army veterinarians and assisting in protocol design. My participation in this project will increase markedly in years 2 and 3.

## Task 4 and 5

As noted above we have developed a highly sensitive methodology capable of identifying genetic variation using a small number of animals. We have also validated this methodology in two separate populations 1) a publicly released dataset consisting of multiple breeds and 2) a quasi-military population of dogs: the retired racing Greyhound.

Besides Task three we have also completed task 5a. Osteosarcoma is the most common cause of death in retired racing Greyhounds, but is extremely rare in show (i.e.; AKC) Greyhounds. We collected samples from Osteosarcoma Positive Racing Greyhounds (n=12) and Osteosarcoma Negative Racing Greyhounds (n=12) in collaboration with the PI. Samples from AKC show dogs (n=12), were collected in collaboration with the AKC at a show in California. Samples were shipped overnight on ice for processing in the laboratory at The Research Institute at Nationwide Children's Hospital. All samples were verified for registration and pedigree information (for racing Greyhounds, NGA registration; for show Greyhounds, AKC registration) and were determined to be unrelated to at least the level of grandparents. Suspected osteosarcomas were confirmed by a Veterinary Oncologist/Pathologist after surgical excision of the tumor. All blood samples were obtained prior to initiation of chemotherapy. For controls, we invited owners of adopted retired racing Greyhounds who were osteosarcoma free and >10 years of age to participate in this research project. Interested owners were screened for inclusion, and subsequently informed consent for blood collection was obtained from the owner, and sample collected by a trained veterinary technician in 1-2 7 mL BD lavender top tubes.

Results of this study are provided in Dr. Alvarez's report. My role in the development of this methodology has been critical. I was able to acquire genetic material and appropriately identify dogs that were ideal for our new methodology. My direct contribution was to aid in the development of appropriate animal protocols and to coordinate with members of the military clinical team at Lackland AFB, as described above. I was involved in meetings with appropriate personnel and aided in the vetting of Mrs. Michelle Perez, our veterinary technician at LAFB. I has also worked diligently in refining the DAPER database and selecting appropriate criteria for study. My primary role in the study has not occurred yet. The majority of my responsibilities will take place in the upcoming years

Task 6- Adaptation of existing resources, data storage and hosting: There have been two areas of progress:

- 1.) Development of a Canine Medical Record System
- 2.) Design of a Workflow for digitizing paper medical records

#### Development of a Canine Medical Record System (CMRS)

A prototype of a research CMRS has been completed. Our research CMRS allows the creation, search, and modification canine medical records based on a group of standard medical record forms used in the military like the Form 1829, Immunization Form, Death Certificate Form, and Master Problem List. The software is meant as a research tool, rather than an operational tool, for viewing, searching, and analyzing medical records that are captured as part of our digitization workflow described below. Our goal is to populate the CMRS database with data generated from our automated digitization process described in the next section. My involvement in this portion of the project has been primarily as an advisor. Details of this portion of the project are given in Dr. Alvarez's report.

#### Digitization Workflow

A preliminary design of a workflow to digitize (i.e. make database searchable) an archive of canine medical records at Lackland. The workflow is designed to be automated and scalable. A critical part of this workflow is incorporating third-party software for OCR (Optical Character Recognition), ICR (Intelligent Character Recognition) and HWR (Handwriting Recognition). A number of various software packages were evaluated and we have settled on a particular software vendor, ABBY Flexi-capture. We looked carefully at Form 1829, and the Chronological Medical Record form to ensure that the third-party software is able to recognize check-boxes and columns and do a reasonable job with handwriting – which it does.

This workflow will also incorporate the Joint Pathology Center Pathology Reports (and Consultation reports). The text will be entered into a data repository that will be further processed by a different pipeline to evaluate the quality of the OCR, ICR, and HWR. The raw data will then be cross-referenced with a controlled vocabulary (e.g. SNOMED, VeNom, ICD- 9) and the matches, or occurrences of controlled terms in the documents, will be stored in an inverted index or specialized database table. The terms can then be correlated against specific canines and stored in a document database. We have not yet evaluated any of the various Controlled Vocabularies.

### Conclusions

Although my involvement in the first year has been limited, we have remarkable accomplishments to show. I am looking forward to additional participation in this project.